

ALEXANDER CHIEN

(951) 892-9896 | alexchien22@ucla.edu | Los Angeles, CA | www.linkedin.com/in/alexander-chien | www.github.com/alchien22

EDUCATION

University of California, Los Angeles (UCLA)

B.S. Computer Science

Los Angeles, CA

Expected June 2026

- GPA: 3.7/4.0
- Selected Coursework: Data Structures & Algorithms, Neural Networks and Deep Learning, Deep Learning for Computer Vision, Foundations of Computer Vision, Probabilistic Models in Genomics, Signal Processing, Introduction to Robotics
- Graduate Coursework: Advanced Neural Networks and Deep Learning, Computational Imaging, Human Factors in AI

RESEARCH EXPERIENCE

Chang's Group at UCLA NLP

Undergraduate Researcher

Los Angeles, CA

June 2024 - Present

- Investigated long-context memory mechanisms for embodied multi-agent systems using Temporary-LoRA, improving knowledge retention in complex VirtualHome simulations
- Generated a synthetic LLaVA-style dataset of 10,000+ multi-action agent tasks with first-person viewpoints using GPT-3.5 Turbo and Unity, enabling multimodal training on grounded agent interactions
- Co-authored a benchmark paper on embodied web agents, accepted to NeurIPS 2025 Dataset & Benchmarks Track (Spotlight)

UCLA StarAI Lab

Undergraduate Researcher

Los Angeles, CA

September 2023 - August 2025

- Reproduced Graphically Structured Diffusion Models experiments for valid Sudoku generation, analyzing masked attention versus explicit logical constraints
- Implemented a sentential decision diagram (SDD)-based constraint filter for GPT-2 decoding, reducing constraint complexity from $O(n^2)$ to $O(n \log n)$ for a 50 K-word vocabulary
- Investigated gradient estimation for discrete expert selection in Mixture-of-Experts models, extending the SIMPLE estimator with multi-sample perturbations to reduce variance while maintaining unbiasedness

Hallucination Detection

Machine Learning Researcher

Los Angeles, CA

March 2025 - June 2025

- Developed a clinically grounded hallucination-detection framework for large language models applied to EHR data, enabling safer medical question answering through retrieval-augmented generation
- Proposed a calibrated confidence metric integrating factuality with a weighted combination of model confidence and retrieval consistency, improving reliability in patient-information retrieval
- Devised experiments to evaluate factuality and trust calibration across medical QA benchmarks, informing ongoing work on multimodal clinical decision support

PUBLICATIONS

- Yining Hong*, Rui Sun*, Bingxuan Li†, Xingcheng Yao†, Maxine Wu†, Alexander Chien†, Da Yin, Ying Nian Wu, Zhecan (James) Wang, Kai-Wei Chang. Embodied Web Agents: Bridging Physical-Digital Realms for Integrated Agent Intelligence. NeurIPS 2025 Dataset & Benchmarks Track (Spotlight).
- * Equal contribution among first authors; † Equal contribution among secondary authors

INDUSTRY EXPERIENCE

A10 Networks

Multimodal Safety Research Intern

San Jose, CA

June 2025 - Present

- Researched multimodal safety alignment frameworks, conducting a comprehensive review of guardrail methods and limitations in visual grounding, latency, and coverage
- Designed and benchmarked a vision-language guardrail model, analyzing trade-offs between performance, efficiency, and robustness to adversarial or distracting visual inputs
- Initiated a project to curate large-scale multimodal safety data from pretraining corpora, laying the groundwork for a peer-reviewed dataset publication

CreatorHub

Software Engineer Intern

Los Angeles, CA

June 2023 - August 2023

- Built and deployed full-stack features for MVP using Flask, MongoDB, and JavaScript, including brand-competition creation and password-recovery workflows
- Designed and validated JSON schemas with MongoDB to ensure data integrity and implemented secure authentication (SHA-256)

PROJECTS

Robotic Manipulation with Robosuite

Reinforcement Learning Researcher

Los Angeles, CA

May 2025 - June 2025

- Developed and evaluated PPO and SAC agents on lift and door manipulation tasks in Robosuite using a simulated Panda arm
- Achieved 90% success with PPO after 10 M steps and 100% success with SAC in under 1 M steps under dense-reward settings, demonstrating SAC's superior sample efficiency

- Diagnosed sparse-reward failures, identifying mode-collapse behaviors and redesigning reward structures for stable policy learning

Novel View Synthesis

Los Angeles, CA

Collaborator and Co-author

February 2024 - March 2024

- Compared three novel view synthesis approaches: light-field rendering, Neural Radiance Fields (NeRFs), and Gaussian Splatting
- Co-authored a technical article analyzing 3D Gaussian Splatting's efficiency relative to NeRF methods (e.g., Mip-NeRF 360)
- Explored trade-offs among spatial fidelity, rendering time, and memory efficiency for real-time 3D reconstruction

LEADERSHIP & OUTREACH

exploretch.la

Los Angeles, CA

Instructor and Mentor

September 2022 - Present

- Designed and taught workshops on computer vision, generative AI, and NLP for 100+ students at exploretch.la's annual outreach expo
- Mentored high-school students from Title I schools in Los Angeles, introducing core ML concepts such as linear models and CNNs

Bruin Spacecraft Group

Los Angeles, CA

Webmaster

October 2023 - September 2024

- Led a team of 6 to redesign and maintain the club's website, improving usability for sponsors and prospective members
- Collaborated with 7 admin teams to ensure accurate project documentation and progress updates

SKILLS

- Programming: Python, C++, C
- Frameworks and Libraries: PyTorch, scikit-learn, Hugging Face Transformers, DeepSpeed
- Data Analysis: NumPy, Pandas, Matplotlib
- Systems/Deployment: Linux, Git, Weights & Biases, Docker
- Research Areas: Natural Language Processing, Retrieval-Augmented Generation, Computer Vision, Multimodal Models (VLMs)